SUPPLEMENTARY MATERIAL
for

"How the forest interacts with the trees: Multiscale shape integration explains global and local processing" by Georgin Jacob & SP Arun

CONTENTS

SECTION S1: STATISTICAL ANALYSES FOR EXPERIMENT 1

The rationale for using linear mixed effects modelling over the traditional ANOVAs is detailed in the main text. Here we describe the basic LMM framework and include detailed statistical reports for each of the analyses described in the main text.

METHODS

Software. We used the R programming language version 3.6.3 with R studio version 1.3.959 for all statistical analyses. We used the *lme4* package (Baayen et al., 2008; Bates et al., 2015) for Linear Mixed Modelling. Since the *lme4* package does not output statistical significance, we used the *lmerTest* package (Kuznetsova et al., 2017) and *Car* package (Fox and Weisberg, 2018). We report the partial eta-squared (η_p^2) as a measure of effect size since it can be compared across experiments (Richardson, 2011; Lakens, 2013). We used the *effectsize* package to calculate partial eta-squared from a Linear Mixed Model. For graphical summaries of the data we used *fitdistrplus, ggplot, mass* and *ggpubr* R packages (Venables and Ripley, 2002; Wickham, 2009; Delignette-Muller and Dutang, 2015).

Linear Mixed effects Model. The basic LMM model for the GDLD pairs was specified in the R environment as Model = lmer(y~Block*ImagePair +(1|Subject), contrast = list(Block = 'contr.sum', ImagePair = 'contr.sum'), data = <data>, REML =False), where y is either the response time or inverse response time, which specifies that the responses are driven by main and interaction effects of Block and Image pair, with Participant as a random intercept factor (which specifies that it introduces unknown random shifts from each participant).

Since the residuals of the LMM models with inverse reaction times as dependent variable were normally distributed in most of the analyses (Distributions of residuals of 9 out of 15 LMM models used in this study are not significantly deviating from normal distribution tested using Kolomogrov-smirnov test), we additionally used an ANOVA (*anova* function in R) to obtain the F-statistic and significance values.

RESULTS

For the GDLD pairs in the global and local blocks, we had data from 16 participants who made 2 responses for each of 147 image pairs, and we are interested in knowing whether responses are systematically different between the global and local blocks.

To investigate the validity of the assumptions underlying the LMM, we fit the LMM model on both RT and 1/RT to GDLD pairs, with blocks (global/local) and image pairs (147 levels) as fixed factors and participants as a random intercept factor. The residual errors of the LMM model are depicted in Figure S1. It can be seen that both the distribution and cumulative distribution of residual error deviate strongly from normal in the case of RT (Figure S1A), whereas the residual errors are much closer to the expected normal distribution for 1/RT based residuals (Figure S1B).

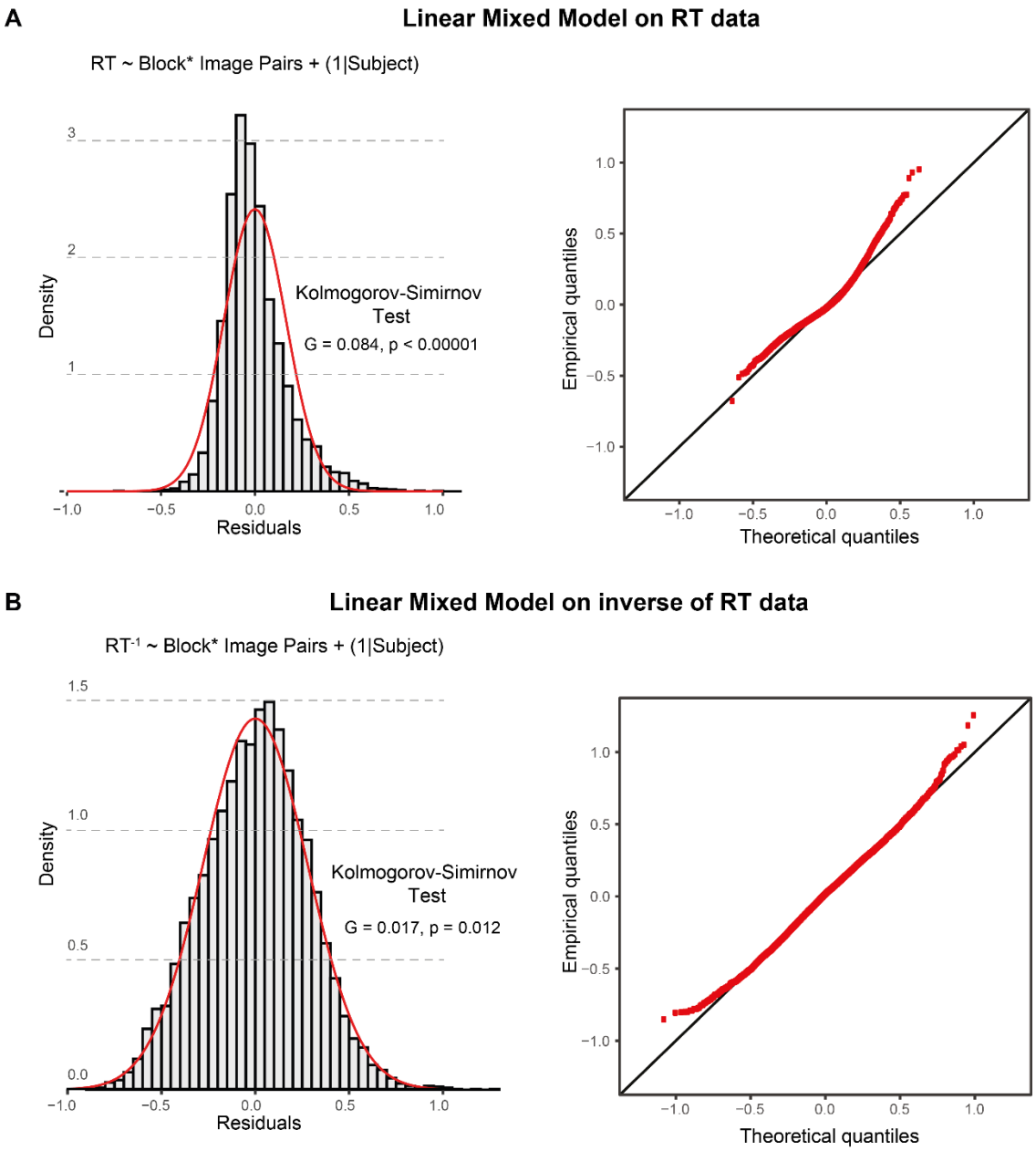

Figure S1. Distribution of residual errors for LMMs using RT and 1/RT

(A) *Left:* Histogram of residuals of an LMM model on RT data (GSLD pairs). This LMM has two fixed factors (blocks and image-pairs) and one random intercept factor (subjects). The red curve shows the normally distribution with the same mean and standard deviation as the residuals. *Right:* QQ plot of the residuals of the observed data plotted against that expected from a normal distribution.

(B) Same as (A) but for the LMM fit to the 1/RT data.

Comparison of statistical test results for RT vs 1/RT ANOVA & LMM models

As an illustrative example, we performed both repeated measures ANOVA as well as linear mixed effects model (LMM) on RT and 1/RT measures in the GDLD pairs in the global and local blocks. It can be seen that the LMM results yield stronger effect sizes with higher statistical significance, since it is based on using raw data, as opposed to the average data used for the repeated measures ANOVA. Further, using

1/RT in the analyses produced stronger effect sizes and higher statistical significance compared to RT-based analyses.

Name of model	Name of effect	Results on RT		Results on 1/RT	
		F-stat	p-value	F-stat	p-value
Repeated measures ANOVA on averages Block*Imgpair + (Subject|(Block*Imgpair)	Block, F(1,15) Image Pairs F(146,2190) Interaction, F(146, 2190)	5.22 2.94 1.62	p < 0.05 p < 0.00005 p < 0.00005	4.36 3.94 1.96	p = 0.0543 p < 0.00005 p < 0.00005
Linear mixed model (Block*Imagepair) + (1|Subject)	Block, F(1,8602) Image Pairs, F(146, 8602) Interaction, F(146, 8602)	124.24 8.28 4.92	p < 0.00005 p < 0.00005 p < 0.00005	**97.75** **7.53** **3.59**	p < 0.00005 p < 0.00005 p < 0.00005

Table S1. Comparison of various statistical models applied on RT & 1/RT. In each case, the F-statistic and p-value is reported for RT and 1/RT data. The linear mixed model on 1/RT was adjudged as the best model (highlighted in **bold**) since its residuals were closest to the theoretically expected normal distribution. In most cases, it also yielded larger F-values and higher statistical significance as well.

Global advantage for GDLD pairs

For GDLD pairs we modelled blocks (global/local) and image pairs (147 levels) as fixed factors with participants as a random intercept factor (see Methods). This revealed a significant main effect of blocks ($F(1,8602) = 97.75$; $p < 0.00005$; $\eta_p^2 = 0.01$) and image pairs ($F(146, 8602) = 7.53$; $p < 0.00005$; $\eta_p^2 = 0.11$) and an interaction between blocks and image pairs ($F(146,8602) = 3.58$; $p < 0.00005$; $\eta_p^2 = 0.06$). A post-hoc analysis revealed that 97 of 147 (66%) image pairs had faster responses in the global block on GSLS pairs, suggesting that the interaction largely modified the magnitude but not the presence of the global advantage effect.

Global advantage for GSLS pairs

For the GSLS pairs, a similar analysis revealed a main effect of block ($F(1,8647) = 413.06$; $p < 0.00005$; $\eta_p^2 = 0.05$) and image pairs ($F(48,8647) = 8.95$; $p<0.00005$; $\eta_p^2 = 0.05$) and an interaction between blocks and image pairs ($F(48,8647) = 6.53$; $p<0.00005$; $\eta_p^2 = 0.03$). A post-hoc analysis revealed that 44 of 49 (90%) image pairs had faster responses in the global block on GDLD pairs, suggesting that the interaction largely modified the magnitude but not the presence of the global advantage effect.

Local-to-global interference for GSLS vs GSLD pairs in global block

Since there is no direct correspondence between the GSLS and GSLD pairs, we performed a linear mixed effects model analysis on inverse response times with interference (GSLS vs GSLD) as a fixed factor and participants as a random intercept factor. This revealed a main effect of interference ($F(1, 8772) = 433.18$; $p<0.00005$; $\eta_p^2 = 0.05$).

Global-to-local interference for GSLS vs GDLS pairs in local block

As before, we performed a linear mixed effects model analysis on inverse response times with interference (GSLS vs GDLS) as a fixed factor and participants as a random intercept factor. This revealed a main effect of interference ($F(1,8564) = 351.16$; $p < 0.00005$; $\eta_p^2 = 0.04$).

Comparing global-local interference and local-global interference

To establish whether the global-to-local interference effect is stronger than the local-to-global interference effect, we compared inverse response times using a linear mixed effects model with block (global/local) and interference (present/absent) as fixed factors and participants as a random intercept factor. This revealed main effects of block (F(1,26015) = 1449.56, p < 0.000005; $\eta_p^2 = 0.05$) and interference (F(1,26015) = 723.08, p < 0.00005; $\eta_p^2 = 0.03$). Importantly, this revealed a significant but relatively weaker interaction effect between block and interference (F(1,26015) = 11.79; p<0.005; $\eta_p^2 = 0.00045$).

Congruence effect for GDLD pairs

To assess the statistical significance of these effects in each task block, we fit a linear mixed effects model on the inverse response times with congruence (2 levels), image pair ($^7C_2$ = 21 levels) as fixed factors and participants as a random intercept factor. This revealed a significant main effect of congruence (F(1,1212) = 36.33; p < 0.00005;$\eta_p^2 = 0.029$ in the global block & F(1,1206) =31.95; p < 0.00005; $\eta_p^2 = 0.026$ in the local block). We also found significant effects of image pair (F(20,1212) =11.8; p < 0.00005; $\eta_p^2 = 0.163$ for global block & F(20,1206) =14.1; p < 0.00005; $\eta_p^2 = 0.189$ for local block). Finally, there was a relatively weak but significant interaction effect between congruence and image pairs in local block (F(20,1212) = 1.13; p > 0.05; $\eta_p^2 = 0.018$ in the global block & F(20, 1206) = 1.79; p < 0.05; $\eta_p^2 = 0.029$ in the local block). A post-hoc analysis revealed that the incongruence effect was present for 18 of the 21 (86%) image pairs in the global block and 16 of 21 (76%) image pairs in the local block. Thus, the interaction largely modified the magnitude but not direction of the incongruence effect.

Congruence effect for GSLS pairs

To assess the statistical significance of the congruence effect in each block, we performed a linear mixed effects model analysis on inverse response times with congruence (2 levels) as fixed factor and participant as random intercept factor. This revealed a significant main effect of congruence (F(1, 4362) =24.39; p < 0.0005; $\eta_p^2 = 0.01$ in the global block & F(1, 4269) = 38.85; p < 0.0005; $\eta_p^2 = 0.009$ in the local block).

Is there a greater effect of stimulus congruence in the global block?

We wondered whether participants showed a larger advantage for the congruent stimuli in the global compared to the local block. To this end, we performed a linear mixed effects model analysis on inverse reaction times to GDLD pairs with block (2 levels), congruence (2 levels), shape pair (21 levels) as fixed-factors and participant as a random intercept factor. This revealed a significant main effect of block (F(1, 2434) =45.75; p < 0.00005; $\eta_p^2 = 0.018$), congruence (F(1, 2434) =66.79; p < 0.00005; $\eta_p^2 = 0.026$) and shape pair (F(20,2434) =21.36; p < 0.00005 ; $\eta_p^2 = 0.149$), block shape pair interaction (F(20, 2434) = 4.26, p < 0.00005 ; $\eta_p^2 = 0.034$) and a congruence- shape pair interaction (F(20, 2434) =1.94, p < 0.05; $\eta_p^2 = 0.016$). Importantly, we observed no significant interaction between block and congruence as would be expected if there was a larger congruent advantage in one block over the other (F(1, 2434) = 0.03; p = 0.86).

For GSLS pairs, we performed a linear mixed effects model analysis on inverse reaction times with block (2 levels), congruence (2 levels) as fixed factors and participants as random intercept factor (we did not include image as a factor because it was unbalanced). This revealed a significant main effect of block ($F(1,8647) = 171.61$; $p < 0.00005$; $\eta_p^2 = 0.019$) and congruence ($F(1,8647) = 58.69$; $p < 0.00005$; $\eta_p^2 = 0.007$). As before we observed no interaction between block and congruence ($F(1,8647) = 0.83$; $p = 0.36$).

We conclude that congruent pairs have an equivalent advantage over incongruent pairs in both global and local task blocks.

SECTION S2: RT CONSISTENCY & RELATIONS BETWEEN BLOCKS (EXPT 1)

Do response times in Experiment 1 vary systematically across image pairs?

The subjects showed a global advantage and incongruence effects in the same-different task but we wondered whether there were any other systematic variations in response times across image pairs. Specifically, we asked whether image pairs that evoked fast responses in one group of subjects would also elicit a fast response in another group of subjects. This was indeed the case: we found a significant correlation between the average response times of the first and second half of all subjects in both the global block ($r = 0.74$, $p < 0.00005$ across 490 pairs; Figure S2) and the local block ($r = 0.75$, $p < 0.000005$ across 490 pairs; Figure S2B). This correlation was present in all four image types as well in both blocks (Figure S2).

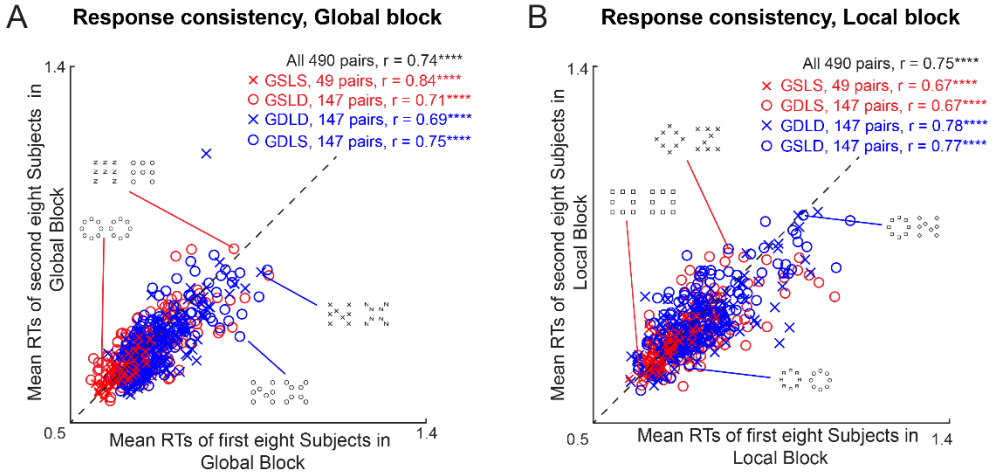

Figure S2. Consistency of response times in the same-different task

(A) Average response times for one half of the subjects in the global block of the same-different task plotted against those of the other half. Asterisks indicate statistical significance (* is $p < 0.05$, ** is $p < 0.005$ etc).

(B) Same as (A) but for the local block.

Are responses in the global and local block related?

Having established that response times are systematic within each block (Section S2), we next investigated how responses in the global and local block are related for the same image pairs presented in both blocks. First, we compared responses to image pairs that elicit identical responses in both blocks. These are the GSLS pairs (which elicit a SAME response in both blocks) and GDLD pairs (that elicit a DIFFERENT response in both blocks). This revealed a positive but not significant correlation between the responses to the GSLS pairs in both blocks ($r = 0.15$, $p = 0.32$ across 49 image pairs; Figure S3A). By contrast the responses to the GDLD pairs, which were many more in number ($n = 147$), showed a significant positive correlation between the global and local blocks ($r = 0.24$, $p < 0.005$; Figure S3A). Second, we compared image pairs that elicited opposite responses in the global and local blocks, namely the GSLD and GDLS pairs. This revealed a significant negative correlation in both cases ($r = -0.20$, $p < 0.05$ for 147 GSLD pairs, $r = -0.23$, $p < 0.0005$ for 147 GDLS pairs; Figure S3B). Thus, image pairs that are hard to categorize as SAME are easier to categorize as DIFFERENT.

Note that in all cases, the correlation between responses in the global and local blocks were relatively small (only $r = \sim 0.2$; Figure S3) compared to the consistency of

the responses within each block (split-half correlation = 0.75 in the global block; 0.74 in the local block; n = 490 & p < 0.00005 for both the conditions; Figure S2). These low correlations suggest that responses in the global and local blocks are qualitatively different.

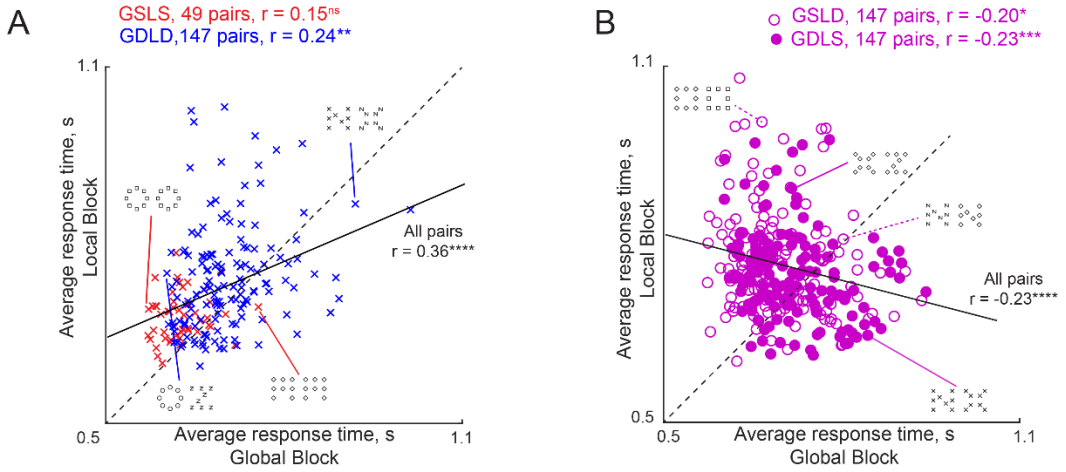

Figure S3. Responses to hierarchical stimuli in global and local blocks.

(A) Average response times in the local block plotted against the global block, for image pairs with identical responses in the global and local blocks. These are the GSLS pairs (red crosses, n = 49) which elicited the "SAME" response in both blocks, and the GDLD pairs (blue crosses, n = 147) which elicited the "DIFFERENT" responses in both blocks.

(B) Average response times in the local block plotted against the global block, for image pairs with opposite responses in the global and local blocks. These are the GSLD pairs (open circles, n = 147) which elicit the "SAME" response in the global block but the "DIFFERENT" response in the local block, and the GDLS pairs (filled circles, n = 147) which likewise elicit opposite responses in the two blocks.

SECTION S3: DISTINCTIVENESS ANALYSES (EXPT 1)

Effect of distinctiveness on same-different responses in the global block

We estimated distinctiveness of a given image as the inverse response time on trials where the image is presented as a GSLS identical image pair. The estimated distinctiveness for the hierarchical stimuli in the global block is depicted in Figure 3A. It can be seen that shapes with a global circle ("O") are more distinctive than shapes containing the global shape "A". In other words, participants responded faster when they saw these shapes.

Having estimated distinctiveness of each image using the GSLS pairs, we asked whether it would predict responses to other pairs. For each image pair containing two different images, we calculated the net distinctiveness as the sum of the distinctiveness of the two individual images. We then plotted the average response times for each GSLD pair (which evoked a "SAME" response) in the global block against the net distinctiveness. This revealed a striking negative correlation ($r = -0.71$, $n = 147$ & $p < 0.00005$; Figure S4A). In other words, participants responded quickly when a given image pair contained distinctive images. We performed a similar analysis for the GDLS and GDLD pairs (which evoke a "DIFFERENT" response). This too revealed a negative correlation ($r = -0.46$, $p < 0.00005$ across 294 GDLS and GDLD pairs; Figure S4B; $r = -0.38$, $n = 147$ & $p < 0.0005$ for GDLS pairs; $r = -0.54$, $n = 147$ & $p < 0.0005$ for GDLD pairs).

If distinctiveness measured from GSLS pairs is so effective in predicting responses to all other pairs, we wondered whether it can also explain the incongruence effect. To do so, we compared the net distinctiveness of congruent pairs with that of the incongruent pairs. Indeed, congruent pairs were more distinctive (average distinctiveness, mean ± sd: 3.31 ± 0.11 s^{-1} for congruent pairs, 3.17 ± 0.14 s^{-1} for incongruent pairs, $p < 0.005$, sign-rank test across 21 image pairs; Figure S4C).

Effect of distinctiveness on same-different responses in the local block

We observed similar trends in the local block. Again, we estimated distinctiveness for each image as the reciprocal of the response time to the GSLS trials in the local block (Figure 3B). It can be seen that shapes containing a local circle were more distinctive compared to shapes containing a local diamond (Figure 3B). Interestingly, the distinctiveness estimated in the local block was uncorrelated with the distinctiveness estimated in the global block ($r = 0.16$, $p = 0.25$).

As with the global block, we obtained a significant negative correlation between the response times for GDLS pairs (which evoked a "SAME" response) and the net distinctiveness ($r = -0.58$, $n = 147$ & $p < 0.00005$; Figure S4D). Likewise, we obtained a significant negative correlation between the response times of GSLD and GDLD pairs (both of which evoke "DIFFERENT" responses in the local block) with net distinctiveness ($r = -0.22$, $p < 0.0005$ across 294 GSLD and GDLD pairs; Figure S4E; $r = -0.24$, $n = 147$ & $p < 0.005$ for GSLD pairs; $r = -0.18$, $n = 147$ & $p < 0.05$ for GDLD pairs). We conclude that distinctive images elicit faster responses.

Finally, we asked whether differences in net distinctiveness can explain the difference between congruent and incongruent pairs. As expected, local distinctiveness was significantly larger for congruent compared to incongruent pairs (average distinctiveness, mean ± sd: 3.08 ± 0.05 s^{-1} for congruent pairs, 2.91 ± 0.11 s^{-1} for incongruent pairs, $p < 0.00005$, sign-rank test across 21 image pairs; Figure S4F).

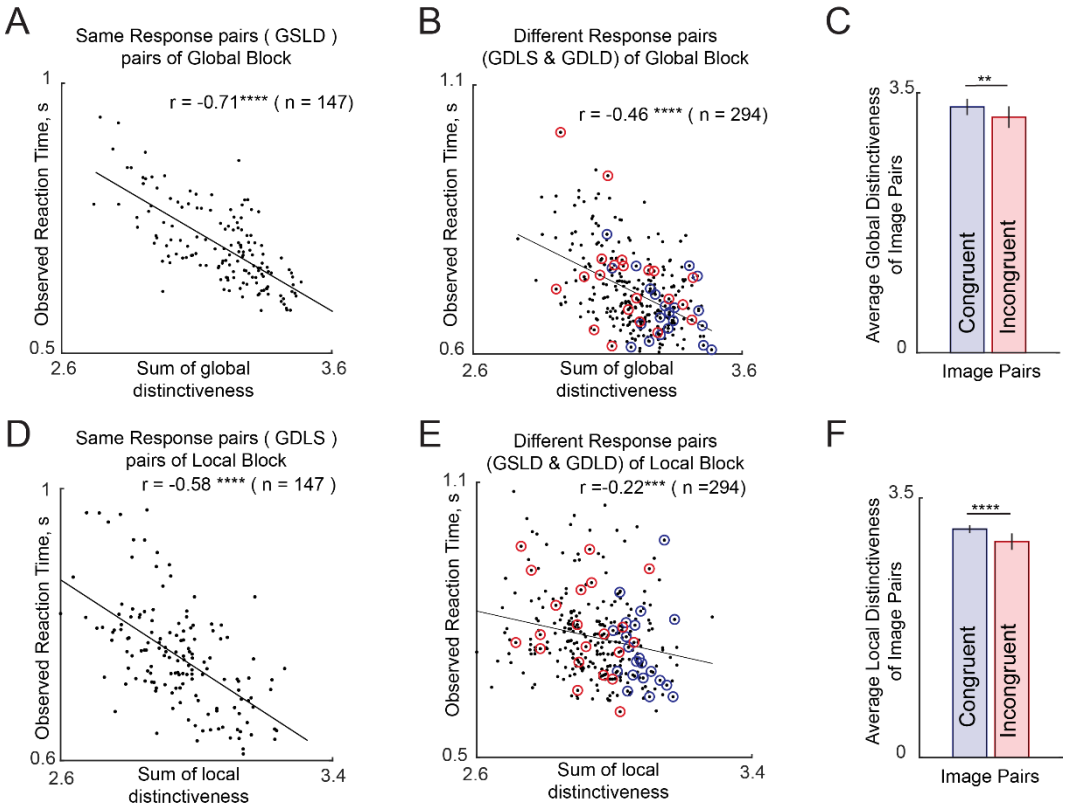

Figure S4. Distinctiveness from GSLS pairs predicts responses to other pairs

(A) Observed response times for GSLD pairs in the global block plotted against the net global distinctiveness estimated from GSLS pairs in the local block.

(B) Observed response times for GDLS and GDLD pairs plotted against net global distinctiveness estimated from panel A. Congruent pairs (red circles) and incongruent pairs (blue circles) are highlighted.

(C) Net global distinctiveness calculated for congruent and incongruent image pairs. Error bars represents standard deviation across pairs.

(D) Observed response times for GDLS pairs in the local block plotted against the net local distinctiveness estimated from GSLS pairs in the local block.

(E) Observed response times for GSLD & GDLD pairs in the local block plotted against the net local distinctiveness estimated as in panel D. Congruent pairs (red circles) and incongruent pairs (blue circles) are highlighted.

(F) Net local distinctiveness calculated for congruent and incongruent image pairs. Error bar represents standard deviation across pairs.

SECTION S4: STATISTICAL ANALYSES FOR EXPERIMENT 2

Global advantage effect in visual search

To establish the statistical significance of the global advantage effect in visual search, we performed a linear mixed effects model analysis on inverse RT with scale of change (global vs local), shape pair (21 levels), common shape (7 levels) as fixed factors and participants as a random intercept factors. This revealed a main effect of scale ($F(1,4696) = 163.24$; $p < 0.00005$; $\eta_p^2 = 0.034$), shape pair ($F(20,4696) = 80.13$; $p < 0.00005$; $\eta_p^2 = 0.254$) and interaction effects for scale & shape pair ($F(20,4696) = 24.38$; $p < 0.00005$; $\eta_p^2 = 0.094$), scale and common shape ($F(6,4696) = 2.41$; $p < 0.05$; $\eta_p^2 = 0.003$) and shape pair and common shape ($F(120,4696) = 1.37$; $p < 0.05$; $\eta_p^2 = 0.034$). There was no main effect of common shape ($F(6,4696) = 0.88$, $p = 0.051$; $\eta_p^2 = 0.001$). A post-hoc analysis revealed that 87 of 147 (59%) of all the matched GDLS-GSLD pairs had a larger average RT for the GDLS pairs, suggesting that these interactions modified the magnitude but not the direction of the effect.

Congruence effect in visual search

To establish the statistical significance of the congruent effect, we performed a linear mixed effects model analysis with congruence, shape pairs as fixed factor and participants as random intercept factor on the inverse reaction times. This analysis revealed a main-effect of congruence ($F(1,664) = 35.87$, $p < 0.00005$; $\eta_p^2 = 0.051$) and shape pair ($F(20,664) = 10.93$, $p < 0.0005$; $\eta_p^2 = 0.248$) and an interaction between congruence and shape pair ($F(20,664) = 2.62$; $p<0.0005$; $\eta_p^2 = 0.073$). A post-hoc analysis revealed that 18 of 21 (86%) of all congruent pairs had faster response times than their corresponding incongruent pairs, suggesting that the interactions modified the magnitude but not the direction of the congruence effect.

Target congruence effect in visual search

To investigate the statistical significance of the congruent target effect, we performed a linear mixed effects model analysis with congruence (2 levels), shape pairs (${}^7C_2 = 21$ levels) as fixed factors and participants as random intercept factor on the inverse of mean reaction times (for each shape pair congruent and incongruent reaction times are estimated by averaging across 5P2 x 2 searches). This analysis revealed a main-effect of shape pair ($F(20,328) = 21.99$, $p< 0.00005$; $\eta_p^2 = 0.573$), and an interaction between distractor congruence and shape pair ($F(20,328) = 6.96$, $p < 0.00005$; $\eta_p^2 = 0.298$). There was no main effect of target congruence ($F(1, 328) = 3.73$, $p = 0.054$; $\eta_p^2 = 0.011$).

Distractor congruence effect in visual search

To establish the statistical significance of the congruent distractor effect, we performed a linear mixed effects model analysis with congruence (2 levels), shape pairs (${}^7C_2 = 21$ levels) as fixed factors and participants as random intercept factor on the inverse of mean reaction times (for each shape pair congruent and incongruent reaction times are estimated by averaging across 5P2 x 2 searches). This analysis revealed a main-effect of distractor congruence ($F(1, 328) = 34.85$, $p< 0.00005$; $\eta_p^2 = 0.096$) and shape pair ($F(20,328) = 18.45$,$p< 0.00005$; $\eta_p^2 = 0.529$), and an interaction between distractor congruence and shape pair ($F(20,328) = 3.28$, $p< 0.00005$ $\eta_p^2 = 0.167$).

SECTION S5: VISUALIZATION OF SEARCH SPACE

We used the reaction times from Experiment-2 to estimate the dissimilarity between shape pairs. Previous studies have shown that 1/RT is a good estimate of dissimilarity between shapes. Multidimensional scaling technique estimates the 2D coordinates of each stimulus such that distances between these coordinates match best with the observed distances. In two dimensions with 49 hierarchical stimuli, there are only 49 x 2 = 98 unknown coordinates that have to match the $^{49}C_2$ = 1,176 observed distances. We emphasize that multidimensional scaling only offers a way to visualize the representation of the hierarchical stimuli at a glance; we did not use the estimated 2D coordinates for any subsequent analysis but rather used the directly observed distances themselves. Two interesting patterns can be seen. First, stimuli with the same global shape clustered together, indicating that these are hard searches. Second, congruent stimuli (i.e. with the same shape at the global and local levels) were further apart compared to incongruent stimuli (with different shapes at the two levels), indicating that searches involving congruent stimuli are easier than incongruent stimuli. These observations concur with the global advantage and incongruence effect described above in visual search.

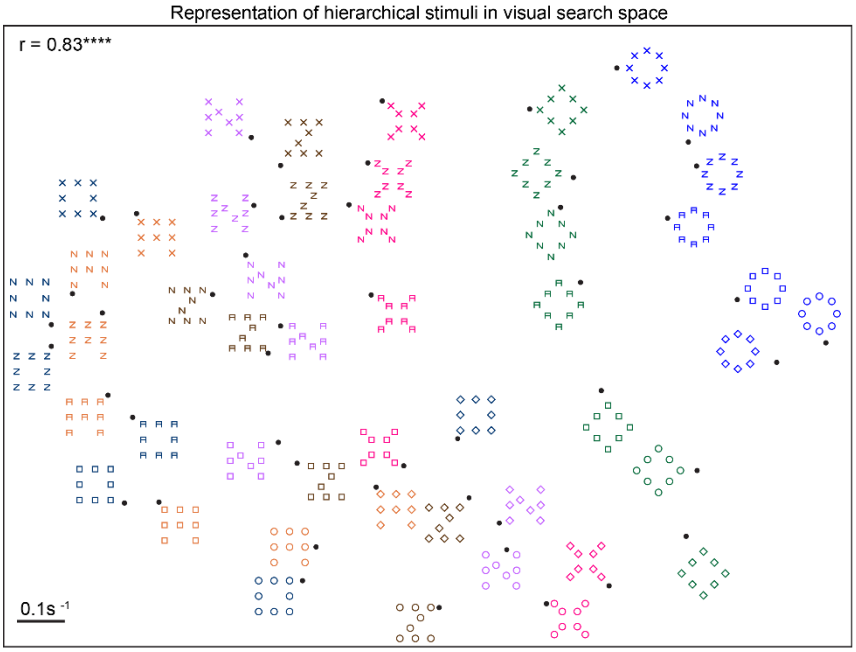

Figure S4. Visualization of hierarchical stimuli in visual search space. Representation of hierarchical stimuli in visual search space, as obtained using multidimensional scaling. Stimuli of the same colour correspond to the same global shape for ease of visualization. The actual stimuli were white shapes on a black background in the actual experiment. In this plot, nearby points represent hard searches. The correlation coefficient at the top right indicates the degree of match between the two-dimensional distances depicted here with the observed search dissimilarities in the experiment. Asterisks indicate statistical significance: **** is $p < 0.00005$.

SECTION S6: COMPARISON WITH OTHER MODELS

Is search for hierarchical stimuli explained better using RT or 1/RT models?

The results in the main text show that search for hierarchical stimuli is best explained using the reciprocal of search time (1/RT), or search dissimilarity. That models based on 1/RT provides a better account than RT-based models was based on our previous findings (Vighneshvel and Arun, 2013; Pramod and Arun, 2014, 2016; Sunder and Arun, 2016). Here we reconfirmed this finding on the visual search experiment in this study (i.e. Experiment 2).

We tested models based on both search response times (RT) and search dissimilarity (1/RT) to identify the best model that accounts for the data. In each case, we fit the full model, in which the net RT or 1/RT corresponding to the search for two hierarchical stimuli is a weighted sum of shape differences at the global and local level as well as cross-scale terms across and within objects (Figure 7B). Because the two models have the same number of free parameters, we compared their quality of fit directly using their overall correlation with the observed data as well as using their residual errors.

Our main finding is that the 1/RT model outperformed the RT model both in predicting the RT and the 1/RT data in terms of correlations (correlations with 1/RT data: 0.88 and 0.81 for the 1/RT and RT models, $p < 0.00005$, Fisher's z-test; correlations with RT data: 0.88 & 0.87 for the 1/RT and RT models, $p = 0.2$). For a finer-grained comparison between the RT and 1/RT models, we compared their residual errors. Here too, the residual error for the 1/RT model was lower than the RT model for both RT & 1/RT data (average absolute error in RT: 0.21 & 0.28 s for the 1/RT and RT models, $p < 0.00005$, rank-sum test across 1176 observations; average absolute error in 1/RT: 0.09 & 0.13 s^{-1} for the 1/RT and RT models, $p < 0.00005$). We conclude that the 1/RT based model provided a better fit to the search data.

Can a simpler multiscale model account for the data?

In the full model described above, the dissimilarity between hierarchical stimuli was taken as a weighted sum of local and global shape differences as well as cross-scale differences both within and across objects. This model yielded excellent fits to the data, but it is possible that a simpler model (using only a subset of these terms) performs just as well.

Comparing the full model with simpler sub-models containing only some types of terms is non-trivial because a complex model will always yield better fits to a given set of data than a simple model by virtue of having more degrees of freedom. Therefore we used a quality of fit measure known as the Akaike's Information Criterion or AICc (Pramod and Arun, 2014, 2016) that penalizes the overall model error by its complexity. The AICc of any model can be calculated as: $AICc = abs\left(N \log\left(\frac{SS}{N}\right) + 2K + \frac{2K(K+1)}{(N-K-1)}\right)$, where N is the number of observations, SS is the sum of squared errors between the model and data across all observations, and K is the number of free parameters in the model. A larger AICc implies a better model.

To compare the quality of fit of two models, we performed a bootstrap analysis. We first resampled the observations with replacement, fit each model and calculated the AICc for each iteration. We then calculated the fraction of bootstrap samples (across 1176 iterations) in which the AICc of one model was larger than that of the other. If this fraction was larger than 95% or smaller than 5% we deemed one model to be superior to the other in terms of the quality of fit.

We fit a number of sub-models that contained various subsets of terms from the full model. Comparing these models on their performance is however not straightforward because some models may have naturally better fits to the data owing to their greater degrees of freedom. We therefore compared the Akaike's Information Criterion or AICc (see above), which takes into account not only the overall residual error between the model predictions and the data, but also penalizes models for having greater degrees of freedom. The results are summarized in Table S1. It can be seen that the full model explains the data better than all sub-models and is superior both in terms of the overall correlation as well as the AICc quality of fit. It can also be seen that global terms contribute the most to the fit, followed by local terms and then by the cross-scale interactions.

Model	***dof***	**Model Correlation**	**Quality of fit AICc** *(mean ± sd)*
G	22	0.67****	3550 ± 44**
L	22	0.45****	3114 ± 38**
X	22	0.34****	2989 ± 46**
W	22	0.30****	2952 ± 42**
GL	43	0.83****	4194 ± 53**
GX	43	0.71****	3619 ± 44**
GW	43	0.71****	3608 ± 44**
LX	43	0.55****	3232 ± 47**
LW	43	0.52****	3175 ± 42**
XW	43	0.39****	2998 ± 45**
GLX	64	0.85*	4298 ± 52*
GLW	64	0.85*	4291 ± 50*
GXW	64	0.74****	3676 ± 44**
LXW	64	0.59****	3250 ± 49**
Full Model (GLXW)	**85**	**0.88**	**4430 ± 52**

Table S1. Comparison of submodels with the full 1/RT model. In each case the 1/RT model containing a subset of the model terms was fit to the full set of 1176 search dissimilarities. The best model, depicted in **bold face**, was the full model containing global (G), local (L), cross-scale across object (X) and cross-scale within object (W) terms. Asterisks in the model correlation column indicate the statistical significance of comparing each model with the best model using a Fisher's z-test on correlation coefficients (* is $p < 0.05$, ** is $p < 0.005$ etc). Asterisks in the AICc column indicate statistical significance of comparing each model with the best model, calculated as the fraction of bootstrap samples in which the AICc was larger than the AICc of the best model.

SECTION S7: SIMPLIFYING HIERARCHICAL STIMULI (EXPT S1)

Having characterized how global and local shape combine in hierarchical stimuli, we wondered whether we can obtain further insights by varying their component properties. One fundamental issue with hierarchical stimuli is that the global shape is formed using the local shapes, making them inextricably linked. We therefore wondered whether hierarchical stimuli can be systematically related to simpler stimuli in which the global and local shape are independent of each other.

These simpler stimuli are shown in Figure S5A. For each hierarchical stimulus, we created an equivalent "interior-exterior" stimulus in which an external contour with the same shape as the global shape encloses a random arrangement of interior elements with the same number and local shape (Figure S5A). We repeated this for two element sizes because the grouping of local elements into a global shape is affected by size (Figure S5B). This design allowed us to ask whether feature integration is similar in hierarchical stimuli compared to the interior-exterior stimuli.

METHODS

Subjects. Eight right-handed human subjects (7 male, aged 23-28 years) participated in the study. All other details were as in Experiment 2.

Stimuli. We designed hierarchical stimuli (H) and matched interior-exterior (IE) stimuli. The hierarchical stimuli were created by combining 5 shapes at the local and global levels in all possible combinations, resulting in a total of 25 stimuli. The interior-exterior (IE) stimuli were derived from the hierarchical stimuli by arranging the local shapes in a fixed configuration, and replacing the global form of the hierarchical stimulus by a solid closed contour (Figure S1A). Shapes were chosen such that their exterior/global version was large enough to accommodate 8 local shapes without intersecting with the local shapes. To investigate how the size of the local elements influences the overall dissimilarity, we created a new set of hierarchical and interior-exterior stimuli in which the local elements were 75% of their original size (size 2; Figure S1B). Thus there were four sets of 25 stimuli used in the experiment (hierarchical and interior-exterior at 2 sizes each). Subjects performed visual search involving all possible pairs of stimuli within each set, with the result that there were ${}^{25}C_2 \times 4 = 1200$ unique searches in the experiment. All other details were identical to Experiment 2.

Model fitting. We fit the multiscale model to all 300 searches corresponding to each stimulus set. Since there were 5 unique parts, there were ${}^{5}C_2 = 10$ model parameters for each group of global, local, across and within terms. Together with a constant term, the multiscale model consisted of 41 free parameters in all. All other details are as in Experiment 2.

RESULTS

Subjects performed visual search using matched hierarchical stimuli and interior-exterior stimuli at two local element sizes (Figure S5A-B). For ease of exposition, we first describe results for the hierarchical and interior-exterior stimuli at the larger size (size 1), and then describe the effect of changing local element size (size 1 vs 2).

Is the representation of interior-exterior stimuli similar to hierarchical stimuli?

Subjects were highly consistent in their performance on searches involving both hierarchical and interior-exterior stimuli (split-half correlation between RT for odd and even subjects: 0.79 & 0.93 for hierarchical and interior-exterior stimuli both of size 1, $p < 0.00005$).

To visualize the underlying representation, we performed a multidimensional scaling analysis on the average search dissimilarities for each set separately. As before, the hierarchical stimuli tended to group according to the global shape (Figure S1C). This trend was even more evident for the interior-exterior stimuli (Figure S5D). Thus, hierarchical and interior-exterior stimuli have qualitatively similar representations.

To quantify these observations, we directly compared the pairwise dissimilarities between hierarchical stimulus pairs and interior-exterior pairs. This revealed a significant positive correlation ($r = 0.65$, $p < 0.0005$; Figure S5E). We note that this correlation is only modest even though the interior-exterior dissimilarities and hierarchical dissimilarities were themselves highly consistent. This implies that there are subtle representational differences between the two sets. We therefore wondered whether the multiscale model would be able to account for these differences. This was indeed the case: multiscale model predictions on both sets were excellent ($r = 0.91$ & 0.96 for hierarchical and interior-exterior sets; Figure S5F). These correlations were virtually the same as the reliability of the data itself ($rc = 0.89 \pm 0.008$ for hierarchical stimuli; $rc = 0.90 \pm 0.003$ for interior-exterior stimuli). Thus, the multiscale model explains nearly all the explainable variance in the data. This in turn implies that whatever subtle representational differences exist between hierarchical and interior-exterior stimuli must arise from systematic differences in their model parameters.

We therefore compared the model parameters for the two sets – and observed several interesting patterns (Figure S5G). First, model terms corresponding to global shape differences were stronger in the interior-exterior stimuli (average magnitude: 0.47 & 0.94 for hierarchical and interior-exterior, $p < 0.005$, sign-rank test on 10 global terms). This is as expected given the stronger clustering by global shape for the interior-exterior stimuli. However the global terms for hierarchical and interior-exterior stimuli were significantly correlated, indicating that the underlying representation is similar ($r = 0.73$, $p = 0.016$). This correlation was even higher across all model terms ($r = 0.85$, $p < 0.00005$ across 41 model terms for hierarchical and interior-exterior stimuli).

Second, model parameters corresponding to local shape differences and cross-scale interactions were weaker in the interior-exterior stimuli (average magnitude of local terms: 0.13 & 0.045 for hierarchical and interior-exterior, $p < 0.005$, sign-rank test; cross-scale across object terms: 0.1 & 0.03, $p < 0.005$; cross-scale within-object: 0.18 & 0.06, $p < 0.005$; Figure S5G). Third, as before, in both sets, model parameters corresponding to local and cross-scale terms were generally correlated with the global terms in the same way as in Experiment 1 (correlation with global terms for hierarchical stimuli across ${}^5C_2 = 10$ shape pairs: $r = 0.9$, $p < 0.005$ for local, $r = 0.86$, $p < 0.005$ for across and $r = -0.57$, $p = 0.08$ for within terms; for interior-exterior stimuli: $r = 0.75$, $p < 0.05$ for local, $r = 0.33$, $p = 0.33$ for across and $r = -0.19$, $p = 0.5$ for within terms). These correlations indicate that model parameters are driven by a common shape representation.

How do model parameters change when local shapes are made smaller?

Next we asked how the estimated model parameters of the hierarchical and interior-exterior stimuli change if the local elements became smaller. In general search times with larger local elements was faster (average search times: 1.72 and 1.90 s for hierarchical stimuli size 1 and 2, $p < 0.005$, sign-rank test on mean response times; 1.32 and 1.37 for interior-exterior stimuli size 1 & 2, $p = 0.83$, sign-rank test). The multiscale model again yielded excellent fits at this size too (correlation between observed & predicted dissimilarity for size 2: $r = 0.93$ for hierarchical stimuli, $r = 0.96$ for interior-exterior stimuli, $p < 0.00005$ in both cases). Importantly, model parameters changed systematically when local elements were smaller (Figure S5H). These changes were similar for both sets of stimuli, suggesting that local element size influences both stimuli similarly. The general pattern is that, when local elements decrease in size, global terms become larger whereas local and cross-scale terms become weaker (Figure S5H).

To summarize, the multiscale model provided excellent fits for searches involving both hierarchical and interior-exterior stimuli even across changes in local element size. Both stimuli were driven by a common underlying shape representation, and their differences were explained by systematic differences in model parameters. The differences in the model parameters indicate that interior-exterior stimuli have more salient exterior shapes with weaker local and cross-scale interactions. The weaker local and cross-scale interactions could be due to the increased salience of the global shape or due to the greater proximity of the local shapes to each other. In subsequent experiments we designed stimuli to distinguish between these possibilities.

How do model parameters change with local shape properties?

The above findings suggest that the multiscale part sum model parameters change systematically with local element size. To further investigate how model parameters change with other local shape properties, we varied the size, position, number and grouping status of the local elements in the interior-exterior stimuli (Sections S2-4). We obtained excellent model fits in all cases, and model parameters varied systematically with these manipulations.

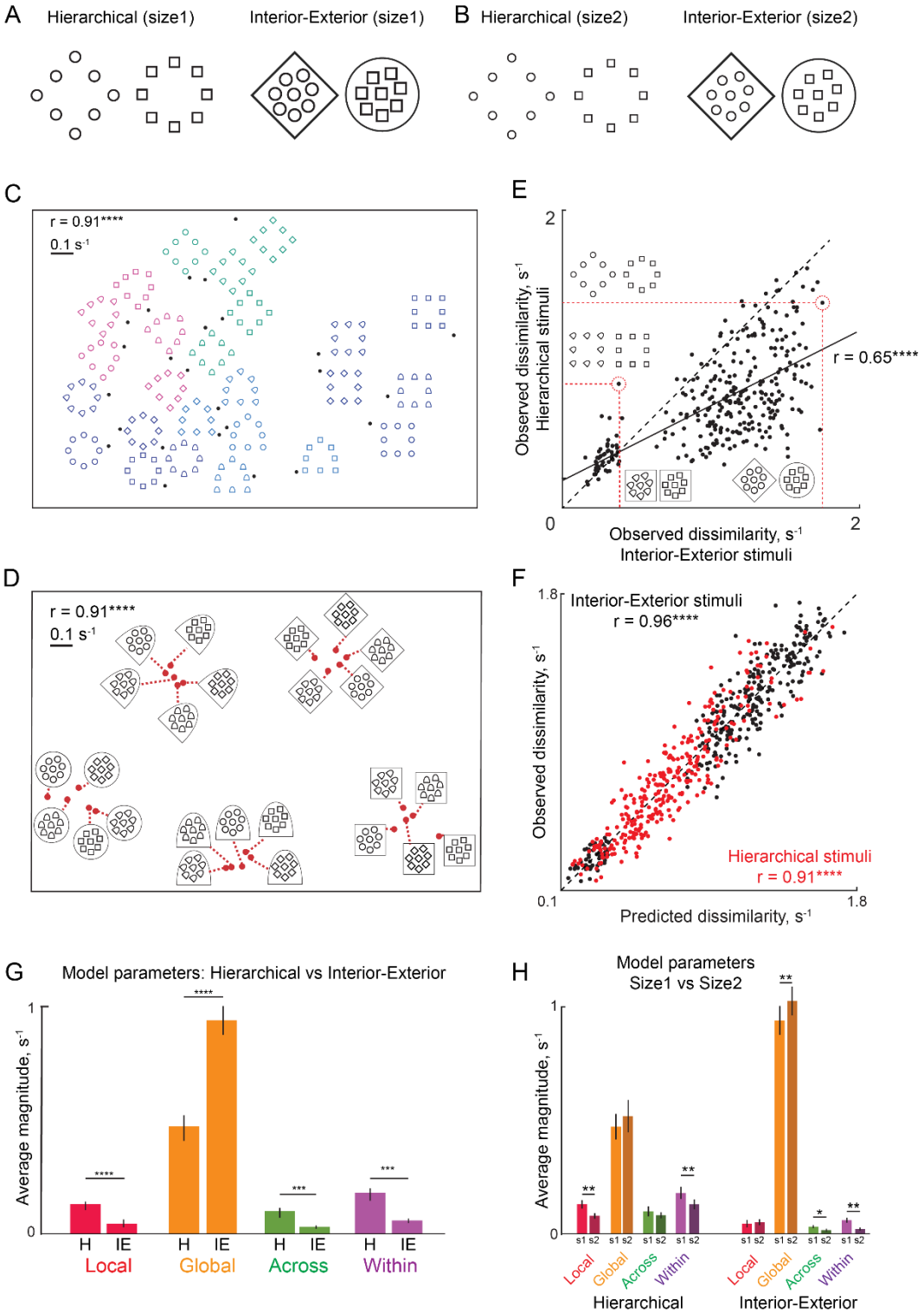

Figure S5. Simplifying hierarchical stimuli into interior-exterior stimuli

(A) Example pair of hierarchical stimuli and interior-exterior stimuli. Understanding global and local shape integration in hierarchical stimuli is complicated by the fact that the global shape is inextricably linked to and formed by the local shape. We attempted to simplify each hierarchical stimulus into an "interior-exterior" stimulus in which the external contour matches the global shape and the shape and number of the internal elements matches the local shape.

(B) Example hierarchical and interior-exterior stimulus pairs with smaller size elements (size 2).
(C) Visualization of the underlying shape representation for hierarchical stimuli (Size 1), as obtained using multidimensional scaling.
(D) Same as (C) but for the matched interior-exterior stimuli (Size 1).
(E) Observed dissimilarities for hierarchical pairs plotted against that of Interior-Exterior pairs (Size 1), with example pairs highlighted (*red dotted lines*). The solid line represents the best-fitting straight line and the dotted line is the y = x line.
(F) Observed versus predicted dissimilarity for hierarchical stimuli (*red*) and interior-exterior stimuli (*black*) for Size 1.
(G) Average magnitude of model terms for hierarchical stimuli (H) and Interior-Exterior (IE) stimuli for Size 1. Note that within-object terms are generally negative but their magnitude is depicted for ease of comparison. Asterisks indicate statistical significance as calculated using a sign-rank test on the model parameters, with conventions as before.
(H) Average magnitude of model parameters for Size 1 vs Size 2 for both hierarchical (H) and interior-exterior (IE) stimuli. Asterisks indicate statistical significance as calculated using a sign-rank test on the model parameters, with conventions as before.

SECTION S8. CHANGING ELEMENT SIZE, POSITION & NUMBER (EXPT S2)

In Experiment 3, we demonstrated that hierarchical and interior-exterior stimuli are driven by a common shape representation. Here we manipulated the position, size and numerosity of local elements in highly simplified interior-exterior stimuli to understand how these changes affect the overall representation.

METHODS

Subjects. Eight right-handed human subjects (5 male, aged 21-28 years) participated in the study. All other details were as in Experiment 1.

Stimuli. We created four sets each containing 25 stimuli. Set 1 was a reference set containing a single exterior shape and a single interior shape (Figure S2). In Set 2, all stimuli were identical to Set 1 except that the interior shape was shifted to the left. In Set 3, the interior shape was double the size of the Set 1 stimuli. In Set 4, there were two local elements of the same size as in Set 1.

Procedure. Subjects performed searches involving all pairwise stimuli within each set. Thus in all there were $^{25}C2 \times 4$ sets = 1200 searches. Subjects performed 98.3 ±0.001% correct trials for each unique search. All details were identical to Experiment 1 except that the Set 1 stimuli measured 3.4° along the longest dimension, and the inter-item spacing was slightly smaller at 3.35°.

RESULTS

We measured visual search performance on four sets of stimuli in which local elements were varied in position, size and number (Figure S2A). Subjects were highly consistent in their search performance across all four sets (split-half correlation between RT of odd- and even-numbered subjects: $r = 0.92, 0.92, 0.87$ & 0.89 for Sets 1-4 respectively, $p < 0.00005$ in all cases). Observed dissimilarity was also highly correlated across sets, indicating that the underlying shape representations are very similar (Figure S2B).

To visualize the underlying shape representation we performed a multidimensional scaling as before. The resulting plot for Set 1 is shown in Figure S2C. It can be seen that stimuli with the same global shape cluster together, indicating that these are hard searches.

As before, the multiscale model yielded excellent fits to the data ($r = 0.95, 0.95, 0.93$ & 0.94 for Sets 1-4 respectively, $p < 0.00005$ in all cases; Figure S2D), implying that variations in the underlying representation across sets due to local element properties are captured by systematic changes in model parameters. These changes are summarized in Figure S2E. The most obvious pattern is that the global terms are substantially larger than all other model terms, indicating that search difficulty is dominated by differences in global shape (Figure S2E). However model parameters varied systematically across the four sets, as discussed below.

We first asked what happens to the shape representation with a change in the local element position (Set 1 vs Set 2). Interestingly, when the local element is shifted away from the centre, local terms became smaller and within-object interactions increased (Figure S2E). However, the underlying shape representation was extremely similar, as evidenced by a strong correlation between the global terms across both sets ($r = 0.96$, $p < 0.00005$).

Next, we analysed how the shape representation changes when the local elements increase in size (Set 1 vs Set 3). We found that global terms decreased, whereas cross-scale interactions both across and within objects increased (Figure S2E). Because some of these changes tend to increase dissimilarity whereas others will cause a decrease, the overall dissimilarity is unlikely to change. Indeed, search times were not systematically different across the two sets (average search times: 0.88 & 0.90 s for Sets 1 & 3 respectively, $p = 0.11$, rank sum test across 300 searches). Thus, increasing local element size increases cross-scale interactions across hierarchical levels.

We then asked how the shape representation changes when the number of local elements increases from 1 to 2. The only significant change was that cross-scale across-object terms were larger in Set 4 compared to Set 1 (Figure S2E). Set 3 is also an interesting comparison with Set 4 because the total area of the local elements is the same in both Sets. Here, we found that global terms were larger, but within-object interactions were smaller for two local elements (Set 4) compared to one large element (Set 3). Taken together, these changes mean that increasing the number of elements increases cross-scale interactions compared to a single small element, but the net increase is still much smaller compared to having a single large element of the same size.

To summarize, visual search for interior-exterior stimuli across changes in local element position, size and number is explained extremely well by the multiscale model. Moving local elements away from the centre (closer to the exterior shape), increasing their number, or increasing size all led to increased cross-scale interactions.

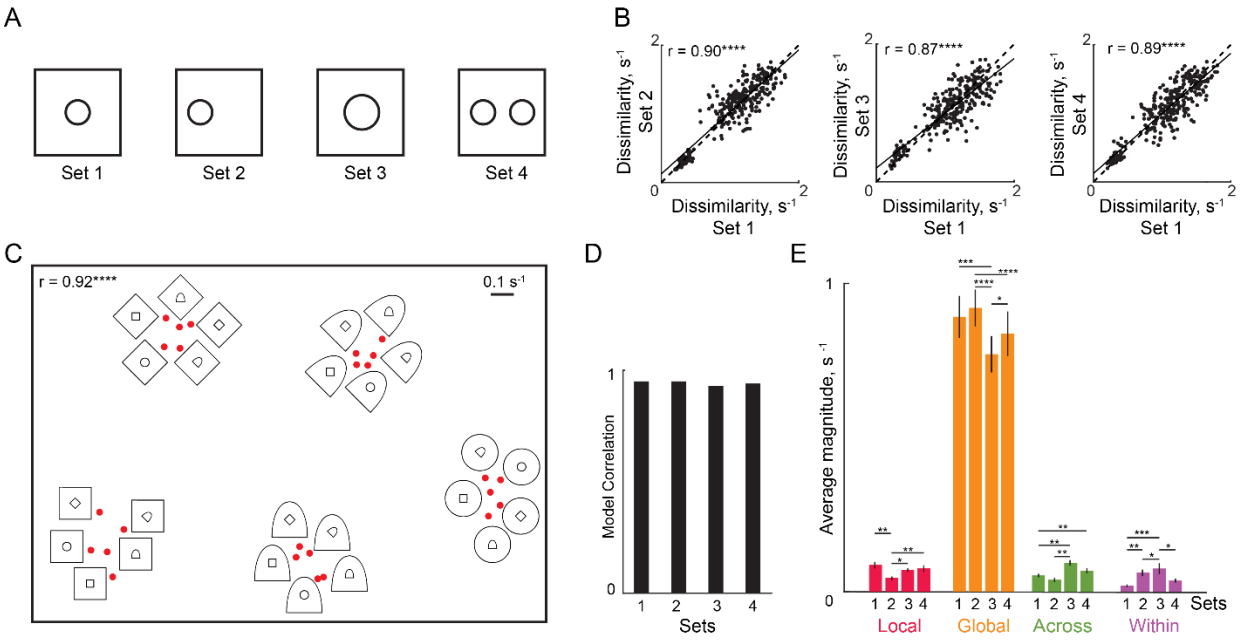

Figure S6: Effect of local element position, size and number

(A) Example stimuli from Sets 1-4. Set 1 is the reference, with the local shape at the centre of the exterior contour. In Set 2 the interior shape is shifted away from the centre. In Set 3, the local shape is doubled in size. In Set 4, two local shapes are placed equidistant from the centre on either side.

(B-D) Observed dissimilarity of all 300 pairs of stimuli in each set plotted against Set 1. The *solid line* is the best-fitting line and the *dotted line* represents the unit line (y = x).

(E) Visualization of the underlying shape representation for the reference set (Set 1), as obtained using multidimensional scaling. All conventions are as before.

(F) Correlation between predicted and observed dissimilarities for each set.

(G) Average magnitude of model parameters for Sets 1-4. Asterisks represent statistical significance assessed using a signed-rank test: * is $p < 0.05$, ** is $p < 0.005$. All comparisons are not significant ($p > 0.05$) unless marked with an asterisk.

SECTION S9: CHANGING ELEMENT POSITION (EXPT S3)

In the previous section, we showed that moving a local element away from the center of an exterior shape tended to increase cross-scale interactions. Here, we explored this issue further by asking what would happen if the local element was moved even further to intersect the exterior shape or even be located outside it.

METHODS

Subjects. Eight right-handed human subjects (6 male, aged 20-26 years) participated in the study. All other details are as in Experiment 1.

Stimuli. We created interior-exterior stimuli with a single local shape whose bounding box was quarter the area of that of the global shape (Figure S7A). We also modified the local shapes to be larger in size compared to the Experiment 3 so as to increase the salience of the local and cross-scale terms and reduce the dominance of the global terms. We created four sets of 25 stimuli each by combining five shapes in the interior with the same five shapes for the exterior contour in all possible ways. The four sets were identical except for the position of the local shape: it could be at the centre (Set 1), between the centre and the left edge (Set 2), centred on the global shape contour (Set 3) and finally located outside the exterior shape (Set 4). These are depicted in Figure S7A. To avoid novel conjunctions with the exterior shapes as the interior shape is moved, we used shapes with vertical edges on both left and right sides, with the result that all shapes differed only in contours on the top or bottom sides. The interior shape was presented in a green colour to facilitate grouping particularly for Set 3 where the interior and exterior contours overlap.

Procedure. Subjects performed an oddball search task on 4 x 4 search arrays as before, with the largest item measuring 4.1°. All other details are same as that of Experiment 2.

RESULTS

In this experiment, subjects performed searches on sets of interior-exterior stimuli in which the center element varied in position. Subjects were highly consistent in their search performance on stimuli in each set (split-half correlation between RT of odd- and even-numbered subjects: $r = 0.82, 0.87, 0.80$ & 0.81 for Sets 1-4 respectively, $p < 0.0005$ in all cases). The observed dissimilarity was also extremely similar across Sets, suggesting that the underlying shape representation is qualitatively similar (Figure S7B). However, search difficulty varied systematically across sets (average search times: 2.00, 1.97, 2.35 and 2.13 s for Sets 1-4), with Set 2 being the easiest ($p < 0.05$, rank-sum test across 300 searches of Set 2 with all other sets) and Set 3 being the hardest ($p < 0.00005$, rank-sum test across 300 searches of Set 3 with all other sets).

To visualize the underlying shape representation, we performed multidimensional scaling as before. In the resulting plot, shown for Set 1 (Figure S7C), it can be seen that stimuli are still clustered according to their global shapes but the grouping is not as strong as in Experiment 3 (i.e. compared to Figure S2C).

As before, the multiscale model yielded excellent fits to the data ($r = 0.93, 0.93, 0.93$ & 0.92 for Sets 1-4 respectively, $p < 0.00005$ in all cases; Figure S7D), implying that variations in the shape representation due to local element position is captured by systematic changes in model parameters. These changes are summarized in

Figure S7E. Unlike the previous experiment (Section S3) where global terms dominated all others, global terms were comparable in magnitude to other model terms and did not vary with element position (Figure S7E). We observed systematic changes in model parameters across sets, as detailed below.

We observed a non-monotonic change in model parameters across sets: local terms became larger from Set 1 to Set 2 as in the previous experiment but became much smaller for Set 3 (where the local & global contours overlap) and increased again from Set 3 to 4 (Figure S7E). Cross-scale across-object terms also followed the same pattern although they did not show as big a drop for Set 3 as the local terms (Figure S7E). Cross-scale within-object interactions were strongest when the local shape was at the centre and decreased in magnitude as its position shifted to the left.

Sets 2 & 4 are an interesting comparison because the local shape is equally far away from the edge of the exterior contour, but different both in terms of being inside vs outside as well as distance from the centre of the exterior contour. Compared to Set 2, local and cross-scale terms (both across and within) were smaller in Set 4 (Figure S7E).

To summarize, visual search for interior-exterior stimuli is explained extremely well by the multiscale model across changes in local element position. Overlaying the local shape on top of the exterior contour (Set 3) strongly reduced the contribution of local terms, indicative of interference due to contour grouping. Local elements enclosed within and near to the exterior contour yielded local and cross-scale terms that were the strongest in magnitude, whereas local elements situated outside the exterior contour yielded weak local and cross-scale terms.

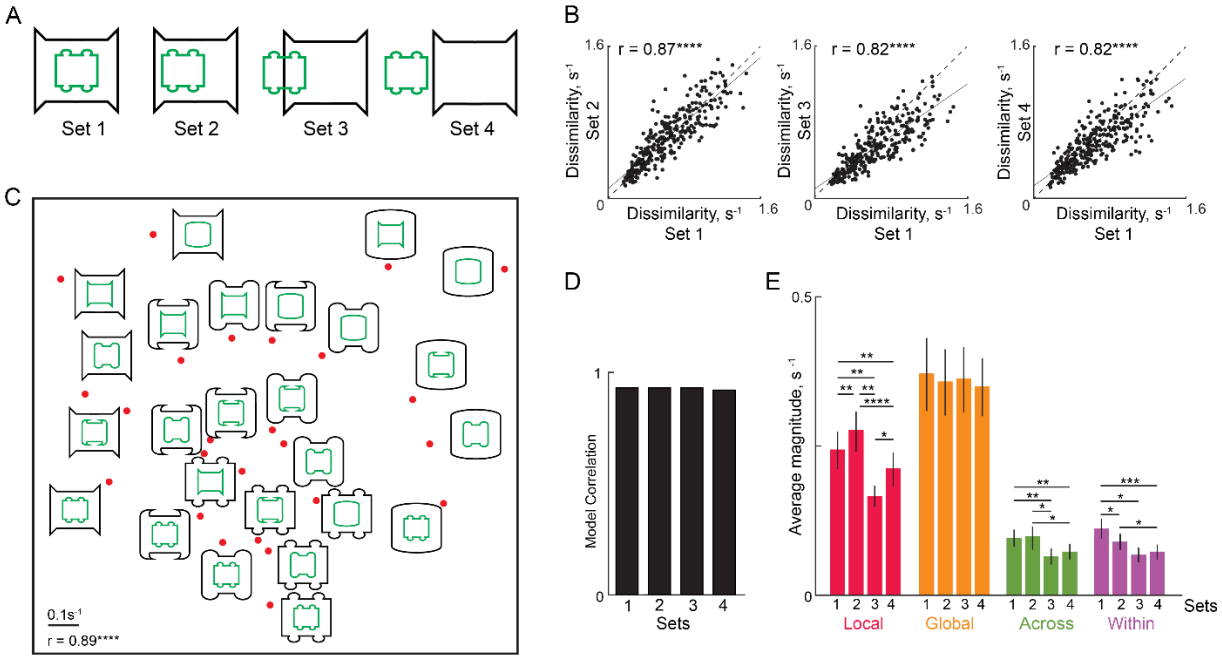

Figure S7: Effect of local element position

(A) Example stimuli from Sets 1-4, in which local elements were made larger compared to the previous experiments and were shifted along a much larger range of positions.

(B-D) Observed dissimilarity of all 300 pairs of stimuli in each set plotted against Set 1.

(E) Visualization of the underlying shape representation for the reference set (Set 1), as obtained using multidimensional scaling. All conventions are as before.

(F) Correlation between predicted and observed dissimilarities for each set.

(G) Average magnitude of model parameters for Sets 1-4. Asterisks represent statistical significance assessed using a signed-rank test: * is $p < 0.05$, ** is $p < 0.005$. All comparisons are not significant ($p > 0.05$) unless marked with an asterisk.

SECTION S10: CHANGING ELEMENT GROUPING

Here, we examine one further influence on the shape representation, namely grouping, by creating stimuli containing identical shapes but differing in their grouping status.

METHODS

Subjects. Eight right-handed human subjects (6 male, aged 20-26 years) participated in the study. All other details are as in Experiment 1.

Stimuli. We created four sets of interior-exterior shapes each containing 25 stimuli. Each stimulus contained four identical interior shapes (Figure S8A). Sets 1 & 2 consisted of stimuli in which the local elements were identical in colour (red in Set 1, green in Set 2). Sets 3 & 4 consisted of stimuli in which two local elements were green and the other two red (Set 3: green along the main diagonal, Set 4: red along the main diagonal). Thus Sets 1-2 have local elements that group by colour and shape whereas Sets 3-4 have local elements that group by shape alone. Half of the subjects performed searches involving Sets 1 & 3 and the other half performed searches involving Sets 2 & 4. In the results, we report the combined results across Sets 1 & 2 as Grouping 1 (G1) and Sets 3 & 4 as Grouping 2 (G2).

Procedure. Subjects performed oddball search exactly as before, with the largest item measuring 4.4°. All other details are identical to Experiment 2.

RESULTS

In this experiment, subjects performed oddball searches for interior-exterior stimuli that either contained local elements of identical colours (G1) or of different colours (G2). Figure S8A illustrates these two types of stimuli. Importantly because the colour and arrangement of the local elements was identical for the target and distractors, these could not serve as cues to guide visual search. Thus differences in search performance across sets can only be due to differences in grouping status.

Subjects were highly consistent in their performance across the two groups (split-half correlation between RT of odd- and even-numbered subjects: $r = 0.79$ & 0.81 for G1 & G2 respectively, $p < 0.00005$ in all cases). Observed dissimilarity was also extremely similar across the two sets, suggesting that the underlying shape representation is qualitatively similar across grouping status (Figure S8C). To visualize the underlying shape representation, we performed a multidimensional scaling as before on the search dissimilarities of G1 pairs. The resulting plot (Figure S8B) shows that stimuli tended to group together by their exterior shape.

For both levels of grouping, the multiscale model yielded excellent fits to the data ($r = 0.93$ & 0.93 for G1 & G2, $p < 0.00005$) indicating that systematic variations across grouping must be captured by systematic variations in model parameters. Indeed, when grouping is disrupted, global and across-object terms increased whereas local and within-object terms decreased (Figure S8D). Thus, in terms of decreasing local & within-object terms, disrupting grouping has the same effect as decreasing local element size (Figure S7H). However, disrupting grouping appears to increase across-object interactions, an effect opposite to that observed with decreased element size (Figure S7H) – this is difficult to reconcile with the other changes.

Taken together, these results show that searches for interior-exterior stimuli are explained extremely well by the multiscale model across changes in the grouping status of local elements, and that grouping tends to make local elements more salient.

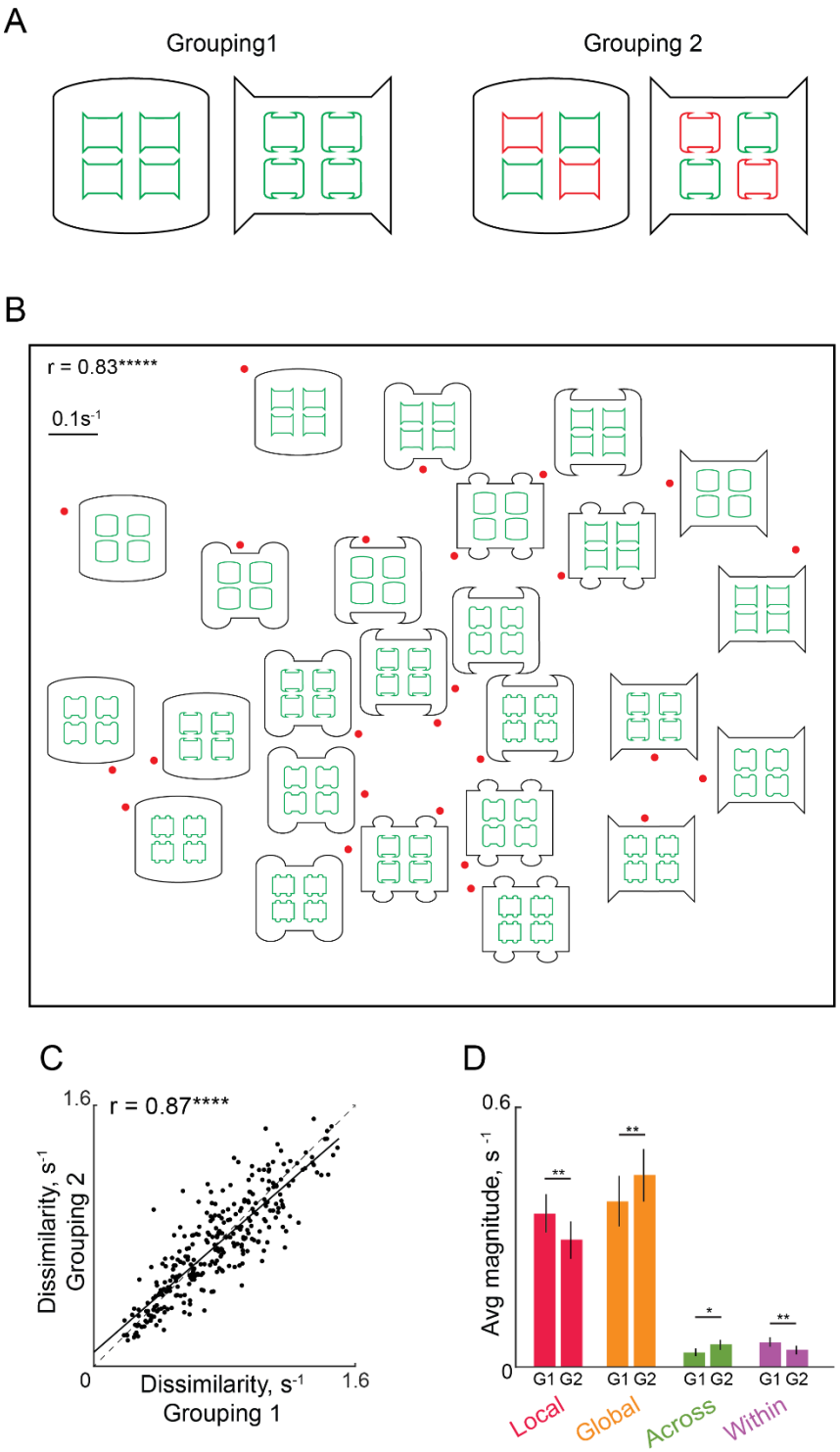

Figure S8: Effect of element grouping on feature integration
(A) Example stimuli from Sets 1&3 representing the two grouping levels (G1 & G2).
(B) Visualization of the underlying shape representation for the reference set (G1), as obtained using multidimensional scaling. All conventions are as before.
(C) Observed dissimilarity plotted against predicted dissimilarity for set G1.

(D) Average magnitude of model terms for sets G1 & G2. Asterisks represent statistical significance assessed using a signed-rank test: * is $p < 0.05$, ** is $p < 0.005$. All comparisons are not significant ($p > 0.05$) unless marked with an asterisk.

SECTION S11. SUPPLEMENTARY REFERENCES

Baayen RH, Davidson DJ, Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. J Mem Lang 59:390–412.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 67:1–48.

Delignette-Muller ML, Dutang C (2015) An R Package for Fitting Distributions. J Stat Softw 64:1–34.

Fox J, Weisberg S (2018) An R Companion to Applied Regression, Third. Thousand Oaks, CA: SAGE Publications Ltd.

Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Softw 82:1–26.

Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 4:863.

Pramod RT, Arun SP (2014) Features in visual search combine linearly. J Vis 14:1–20.

Pramod RT, Arun SP (2016) Object attributes combine additively in visual search. J Vis 16:8.

Richardson JTE (2011) Eta squared and partial eta squared as measures of effect size in educational research. Educ Res Rev 6:135–147.

Sunder S, Arun SP (2016) Look before you seek: Preview adds a fixed benefit to all searches. J Vis 16:3.

Venables WN, Ripley BD (2002) Modern Applied Statistics with S, Fourth Edi. New York: Springer.

Vighneshvel T, Arun SP (2013) Does linear separability really matter? Complex visual search is explained by simple search. J Vis 13:1–24.

Wickham H (2009) ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer New York.